EXPLORATION AND OPTIMIZATION OF USER EXPERIENCE IN VIEWING VIDEOS ON A MOBILE PHONE

WEI SONG

Faculty of Science and Technology, Queensland University of Technology
Level 8, 126 Margaret Street, Brisbane, QLD 4001, Australia
songw2@qut.edu.au

DIAN TJONDRONEGORO

Faculty of Science and Technology, Queensland University of Technology
Level 8, 126 Margaret Street, Brisbane, QLD 4001, Australia
dian@qut.edu.au

MICHAEL DOCHERTY

Faculty of Science and Technology, Queensland University of Technology
O Block Podium Level 4 423, 2 George Street, Brisbane, QLD 4001, Australia
m.docherty@qut.edu.au

Received (Day Month Year)
Revised (Day Month Year)
Accepted (5 Aug 2010)

Compared with viewing videos on PCs or TVs, mobile users have different experiences in viewing videos on a mobile phone due to different device features such as screen size and distinct usage contexts. To understand how mobile user's viewing experience is impacted, we conducted a field user study with 42 participants in two typical usage contexts using a custom-designed iPhone application. With user's acceptance of mobile video quality as the index, the study addresses four influence aspects of user experiences, including context, content type, encoding parameters and user profiles. Accompanying the quantitative method (acceptance assessment), we used a qualitative interview method to obtain deeper understanding of a user's assessment criteria and to support the quantitative results from a user's perspective. Based on the results from data analysis, we advocate two user-driven strategies to adaptively provide an acceptable quality and to predict a good user experience respectively. There are two main contributions from this paper. Firstly, the field user study allows a consideration of more influencing factors into the research on user experience of mobile video. And these influences are further demonstrated by user's opinions. Secondly, the proposed strategies – user-driven acceptance threshold adaptation and user experience prediction – will be valuable in mobile video delivery for optimizing user experience.

Keywords: User experience; mobile video; acceptance threshold; acceptance grade; acceptable degree; user experience prediction.

1. Introduction

The advance of multimedia and mobile techniques has boosted user's demand for viewing videos on their mobile phones, anytime and anywhere. As users' satisfaction decides if they will use a product, providing users an optimal viewing experience is becoming compelling. However, user experience is a consequence of a user's internal state, the system and the context within which the interaction occurs [1]. Simply studying the characteristics of a system (e.g. complexity, usability, functionality, *etc*.) cannot provide a thorough understanding about user experience or a proper method to improve it. In mobile video usage, to ensure the provided video is not isolated from end-users, many studies have been conducted in respect of user-perceived quality [2-4] and acceptability or acceptance of the mobile video [5-7].

The evaluation of user-perceived quality, also called perceptual, affective quality, is commonly carried out with standardization approaches of subjective video quality assessment, such as ITU's recommendation [8]. In these approaches, test subjects are asked to use assigned scales (5/11) to mark the test videos that have various compression qualities or transmission qualities. However, some researchers have pointed out that these assessments lack ecological validity and may not predict whether a particular quality level is acceptable [9]. McCarthy and Knoche et al. [5, 9, 10] suggested an acceptability approach, which is efficient in measuring the minimum acceptable quality for meeting user's expectation. However, this approach is unable to discriminate outside of the acceptable threshold levels and to accurately compare a pair of videos with similar quality [7]. Jumisko-Pyykkö et al. [7] recommended the study of satisfaction of quality parallel to acceptability. Nevertheless, it is questionable if it is useful to evaluate the satisfaction of an unacceptable video.

Although the above approaches take into account user's perception and requirements with objective video quality, it is insufficient to study user experience because user experience is influenced by many factors, such as content, context, motivation, expectations, and user profiles [11-14]. Unfortunately, most studies were conducted in a laboratory environment, which suffer from limited realism. Only a few studies noticed the impact of usage context on the evaluated quality of mobile TV [15, 16]. Users may pay less attention to the video content when they have another task because they worry about the risk of accidents and lapses [15]. Also, users' entertainment of viewing videos on mobile phone is influenced by context [16]. As to the impact of user profiles, it was found that people's experience in image/video processing somehow impacted user-perceived quality [17]. However, this research is limited in using one user profile and not considering mobile video scenarios. Due to the important influence of usage context and user profiles on user experience, it is necessary to conduct a user study with real or potential users in real usage contexts so as to gain accurate user experience.

Concerning research methods underpinning the studies on user experience, both quantitative methods and qualitative methods are commonly used. The subjective assessments on perceived quality or acceptability generally belong to the quantitative method employing numerical scores to ascertain the relative acceptance level of a mobile video. However, the principles behind the quantitative scores placed by test subjects are rarely studied [18]. Qualitative methods are helpful to interpret and understand how user experience is impacted [11-14, 18]. Whereas owing to the non-measurable variables, it is difficult to directly use the understanding from the qualitative study into the user-driven delivery of mobile video.

This paper aims to better understand how user's viewing experience of mobile video is influenced by content, context, video encoding parameters, and user profiles, and based on their relationships to develop proper strategies for optimizing user experience. We conduct a field user study with potential and real mobile video users under two typical usage contexts: relaxed scenario and nervous scenario. Quantitative methods are used to gain the acceptance threshold and acceptable degree for news and sport videos encoded with various encoding conditions. The acceptance threshold is used to indicate the bottom line of the acceptable quality and the 10-scale acceptable degree is used to evaluate the acceptable level of video quality. Simultaneously, qualitative methods are adopted to define a user's evaluation criteria and to interpret and verify the results from quantitative analysis. Finally, in light of the findings from the quantitative and qualitative studies, we present a user-driven acceptance threshold adaptation and computational model for the prediction of a good user experience.

The following content is organized as below. Section 2 is the related work. The design details of this user study are described in Section 3. Section 4 presents the results of statistical analysis. Section 5 provides the results derived from qualitative data and discusses the relativity of subjective opinions with the results of the quantitative study. Two user experience strategies are addressed in Section 6, followed by conclusions and future work in Section 7.

2. Related Work

A typical problem in mobile video study is how to optimize user experience. Studies often consider the strict technical or resource constraints, such as limited bandwidth, processing and displaying capability of mobile devices, but ignore other key elements of mobile video, such as complex usage contexts and diverse user needs.

2.1. *Influencing effects of UX*

In order to optimize user experience, the first step is to understand what influences a user's experience [14, 15]. Many subjective video quality assessments have found user's perceptual quality is significantly influenced by technical factors, such as, spatial and temporal resolution, bit rate, and content features [4, 10, 19, 20]. And their correlations have been used to improve the perceptual video quality [4, 21-23]. However, a user's perception of the quality of mobile video is only part of user experience. Previous user studies on mobile TV claimed user's viewing experience is influenced by social-psychological effects, such as consumption model, service, context, user profile, and their motivations [12-14, 24-27]. Among these factors, the consumption model and the service should be considered by service providers; while other factors have to be considered in a user-centered quality study.

Generally, researchers tend to distinguish users by age, work, education background or their relationship with modern technologies. According to Orgad [11], adults aged 18-34 will be the primary mobile TV users who are familiar with technologies build into mobile devices and interested in new technologies. Also teenagers and children are likely to form another important user group who may think mobile video is fashionable, and businessman may use it during travel or work breaks. Some studies have revealed that different user types have different attitudes to mobile video and different preferred content types [28]. However, no research gives an answer as to whether there is a relationship between the user profile and a user's judgment to the video quality.

The main motivations of using mobile video can be summarized as killing time, being up to date with news, relaxing, sharing and entertainment [27-30]. Since users' motivations determine when and where they use it, when people are traveling or waiting for transportation or friends, they would like to consume the time with mobile video; also, during the lunch/coffee time or work break, people watch mobile videos for relaxing or sharing information with other people [12, 28, 29, 31]. Although it has been found that different results may be obtained in the lab and in the field [6], studies seldom addressed the impact of context on people's entertainment with mobile video [16].

Overall, the social-psychological influencing factors (e.g., context and user profile) are important for user experience, but they are non-measurable and uncontrollable. Thus, these factors should be studied with technical factors (e.g., encoding parameters, or transmission parameters) together, so that their impacts on user experience can be compensated or adapted by adjusting the technical factors. In this paper, we analyze multiple influencing factors of user experience, including encoding parameters, content and context, and several user profiles (age, gender, preference, and experience of mobile video). Moreover, we propose user-based strategies to adapt to the non-technical effects of content, context and user profile with measurable factors of encoding parameters.

2.2. *Approaches to user study*

For different purposes, the user study of mobile video can be conducted with quantitative methods and qualitative methods. Quantitative methods, such as surveys [32] and quality assessments, aim to obtain precise measurements and statistical power; while qualitative methods such as interviews [14, 33] and observations [34] attempt to find an in-depth understanding of user behavior and the reasons that govern such behavior in a particular research setting. The two kinds of methods can be used together. For example, Knoche and McCarthy [26] studied the user requirements of mobile TV with both methods. Jumisko-Pyykkö *et al*. [18] applied a qualitative research method into a study of overall audiovisual quality. And the qualitative results supported their quantitative results of evaluation criteria. Generally speaking, the joint use of the two methods is beneficial in considering measurability and interpretability at the same time and verifying each other's conclusions. Unfortunately, the two methods are not well combined to use in mobile video studies.

Of the quantitative methods, the most widely used method is subjective quality assessment such as ACR with Mean Opinion Scores (MOS). This is recommended by ITU-T and first used for television image quality assessment with 5/9/11-point scale [35]. However, some researchers pointed out there are two main limitations of the ITU recommended approach [9]. Firstly, it is hard to map people's perceptions onto the 5 point scale with labels of Excellent, Good, Fair, Poor, and Bad. Secondly, there is a lack of ecological validity to assess new multimedia services due to "the short duration of test material, the absence of a task, and assessing in isolation neglected interaction effects". To overcome these limitations, some new evaluation methods were developed in terms of information assimilation [36], satisfaction [36] and acceptability [6, 10]. The information assimilation indicates the understanding of content semantics, whereas it is more likely impacted by complexity of content and people's comprehensions about certain contents. The idea of acceptability is to identify the lowest acceptable quality level, which is adopted in a mobile video study by McCarthy *et al*. [10] and continually used in their following works [5, 6, 26]. Jumisko-Pyykkö *et al*. [7] applied the same idea, so-called "Acceptance Threshold", to the quality evaluation under transmission errors. However, they argued acceptability is unable to discern a pair of videos with close quality, especially the videos near the acceptable threshold; and they suggested to use satisfaction of quality synchronously [7]. The satisfaction is measured with a scale 0–5/10 [16, 36]. However, it may be meaningless for an unacceptable video.

To conduct a user study, a lab experiment is controllable and replicable in an artificial setting but has low ecological validity. Researchers cannot observe participants in their natural setting, and do not take into consideration people's thoughts and feelings *etc*. The mobility feature of mobile video, however, decides that the watching behavior always happens in a complicated context. Therefore, lab experiments may be inappropriate for studying mobile video user's experience. Field user study is conducted in user's environment rather than a laboratory, and users perform their normal activities rather than contrived tasks. Compared with lab experiments, field use study is more suitable for the study of user experience on mobile video as more authentic data can be obtained.

Weighing all the above pros and cons, we have undertaken our user study in typical usage contexts. Also, we recruit participants from the expected user population, create realistic tasks and use real material [9]. Both quantitative and qualitative methods are involved in this study because the two methods can be conducted simultaneously and reinforce one another in functions. With regard to the assessment approach for mobile video quality, this paper employs "acceptance threshold" to indicate the minimum acceptable quality level for mobile video users; and "acceptable degree" (10 scales) to denote how good a user feels an acceptable video. In addition, "acceptance grade" (Fair, Good, Very Good, Excellent) is adopted to map the score of acceptable degree into a describable and accessible user experience grade.

All details of the user study are given in the next section.

3. Field User Study

To ensure realism of user experience data, this user study is carefully designed. The following subsections describe the selection of test content, test context and participants; and present the test procedure, assessment method, and data analysis methods.

3.1.1. *Test materials*

We selected two types of video content, news and sports, as the source of test materials. The selection is based on their popularity for mobile television [27, 28, 37] and their representative characteristics. News video provides the most amount of information and is updated very fast. It mainly contains head and shoulder sequences. While sports video is characterized by fast movement and small objects, so that its quality is often a source of complaint by viewers, especially when displayed on a small screen. Since it is important to use real materials for user-oriented study [9], we recorded the news video from Channel Ten News in Australia and the live football match between America versus Brazil in FIFA 2009. The two recorded video sources had a high quality of 1280x720 resolution and 8542kbps bit rate in MPEG2 format. Each of the video sources was then segmented into meaningful 20 short clips based on the semantics. For the news video, each clip consists of the sequences of a reporter and the sequences of related news events; and for the football video, each clip is composed of mostly far-away play scenes and a few close-up player scenes. The duration of these clips is about 30–50 seconds.

To examine the relationships of objective parameters of video coding with user experience, we transcoded the 20 video clips into the format of H.264 (baseline 3.0) with 20 (2x2x5) different combinations of the following coding parameters:

- two spatial solutions (SR) : s1–320x240 and s2–480x320
- two frame rate (FR) : f1–12.5fps and f2–25fps
- five bit rate (BR) : b1–96kbps, b2–128kbps, b3–256kbps, b4–512kbps, and b5–768kbps

The combinations are listed in Table 1.

Table 1. Encoding parameter combinations for each content type

No.	Parameters	No.	Parameters	No.	Parameters	No.	Parameters
1	b1,s2,f1	6	b5,s2,f2	11	b1,s1,f2	16	b5,s1,f1
2	b2,s2,f1	7	b4,s2,f2	12	b2,s1,f2	17	b4,s1,f1
3	b3,s2,f1	8	b3,s2,f2	13	b3,s1,f2	18	b3,s1,f1
4	b4,s2,f1	9	b2,s2,f2	14	b4,s1,f2	19	b2,s1,f1
5	b5,s2,f1	10	b1,s2,f2	15	b5,s1,f2	20	b1,s1,f1

Since the video sources had a very high quality, they would not impact the experimental results when they were encoded into the quality-impaired test materials. In addition, though audio quality was not studied in the paper, we kept the audio to be encoded with the same format: AAC3, 64kbps and 48kHz to help participants understand the contents. Ultimately, we used total 40 video clips with audio (20 for news and 20 for sport) as the test materials.

3.1.2. *Test tools*

Due to the popularity and good functionality of displaying videos, an iPhone 3GS was chosen as the test tool, which has a 3.5-inch multi-touch screen with 480x320 pixels resolution 163 ppi. In order to make the test convenient and give participants more freedom to complete the test, we specifically developed an ad-hoc iPhone application for conducting the test. The user interface is shown in Fig 1. Screenshot 1 shows the video contents list. The button of viewing test results and the button of adding a tester are available on the left and right side respectively at the top of this page. Before starting the test, participant's information is collected in screenshot 2. The required information includes name, age, gender, test location, whether or not having experience in mobile video, and whether or not they liked the tested video contents. When choosing certain content by touching the corresponding cell from screenshot 1, test video clips for the chosen content will be displayed on the screen one by one (examples see screenshot 5 for news and screenshot 6 for sports). Each clip follows an assessment page shown in screenshot 3. To evaluate the acceptability of a video clip, two buttons—"Unacceptable" and "Acceptable"— are clickable. If the "Acceptable" button is clicked, a slider will appear to allow participants scoring the acceptable degree by dragging it to a corresponding value (1-Fair to 10-Excellent). Then pushing the "Next" button will start to play next test clip. Eventually, all the user information and scores are stored into a file and are able to be viewed from a results page in screenshot 4.

Fig 1. Screenshots of test iPhone application

3.1.3. *Test Context*

According to the studies on mobile TV [27, 28, 30], the main purposes of watching mobile videos are to kill time, to keep up-to-date with news, to entertain, and others. Correspondingly, the typical usage context of viewing mobile video include: spending time in vehicles; waiting at bus/train stations or lounges; lunch/work/class break; at home; and so on. Despite diverse locations and motivations, we can still generally classify the usage contexts into two categories: relaxed scenario and ‘nervous’ scenario. The former refers to the scenario that people are free to watch a video and they can concentrate on the video content although sometimes the surroundings are noisy (e.g., lunch break or class break); the latter refers to the scenario that people take another task into their minds when they are watching a video so that they cannot pay attention to the video (e.g., taking a bus and waiting for a bus). For simulating the real usage contexts, we defined two situations to conduct the user study. The first was called “Campus”. It was at a lounge outside of the university library where students usually sit to take a class break or have their lunch, so it can represent the relaxed scenario. The second was called “Bus”. It was on the shuttle bus commuting between two campuses of a university, which represents the nervous scenario. There is another usage context — watching mobile videos at home — which is outside the scope of this study. It could be considered as a relaxed scenario in a quiet environment that is similar to a lab context. Laboratory experiments have been conducted in many studies [10, 20].

3.1.4. *Participants*

Since the primary users of mobile video are young people [11, 28], our target participants were mainly university students and some young staff. In order to assure the collected data is as real as possible, we recruited participants on-site in the predefined two locations, rather than assigned the recruited participants to complete the test with the given tasks [16]. First of all, we observed the people at the lounge area or the bus stop to select the potential responders based on the population representativeness of the sample (such as different gender and age). Then, we asked them to help us do the study and earn a coffee voucher for their time. In this way, 32 participants were recruited, stratified by age (17–40), gender (22 females and 20 males), experience in viewing videos on mobile phone (18 have no experience and 24 have experience), and preference for content types (like, neutral and dislike; the number of people differs for different types). All participants stated they had normal vision.

The way of recruiting participants in field has two advantages. On the one hand, the responder’s viewing experience is close to a real user’s experience. For example, the responder on the bus is a real passenger, and does have another task (e.g., attention to the destination) in his/her mind. Sometimes we can even pick up the person who is watching something on an iPhone. On the other hand, variety of samples can be guaranteed by selecting different gender, age, and culture. However, the variety is also a limitation because the criterion of assessment varies from person to person so that the collected scores may be complicated. In spite of this, our study tries to discern as many influencing factors as possible.

3.1.5. *Procedure and assessment method*

The test procedure was performed through the custom-designed iPhone application. To begin with, participants' profile information was collected by entering the information (e.g., name and age) or selecting an answer (e.g., Have you watched videos on your mobile phone). We then asked participants to evaluate several demo videos (including low and high quality videos) for the purpose of making them familiar with the test process and getting their own assessment criteria built. Then we told them to watch the videos in a comfortable position and let them do the test alone. We also tried to keep a distance with participants and not to disrupt their watching. It is reasonable to consider these details because in a previous user study we observed that participants often felt nervous and hesitated to make a decision if the tester watched their assessment process. We believe this will affect participants to rate correctly, and what's worse, it will impact the realism of user experience. All the above measures aim to provide participants a possible real usage context so as to obtain as real as possible of user experience..

Participants might watch one or two content types (news and sports), depending on if they had time or would like to watch. For each content type, 20 encoded test clips were displayed and evaluated one after another in chronological order. It took around 10 minutes to complete one content (including the viewing time and the evaluation time). And all evaluations were automatically saved into a file on the iPhone.

Following the acceptance assessment, semi-structured interviews were performed with 10 participants. Voice recording or field notes were used to record the responses. We would have a glance at their scores first, and based on their provided information asked the questions about the following aspects.

- How do they assess their experience of video quality?
- How do they feel the environment impacts on their viewing experience?

With regard to the quality assessment method, we adopt acceptance assessment in this study. Participants were requested to evaluate two things: (i) whether the video quality is acceptable or unacceptable; and (ii) if acceptable, to what degree it is acceptable. The "unacceptable" is recorded as 0 and the "acceptable" is recorded as 10 scales (1-10), which correspond to four grades: Fair, Good, Very Good, and Excellent. For the sake of clarity, we defined the terms of the assessment indexes used in this paper as follows:

- *Acceptance threshold: 80% "acceptable" ratio of the total assessments or 20% "unacceptable" ratio of total assessments.* That is, if over 80% participants mark a particular video clip as "acceptable", the clip will be acceptable; otherwise, it will be unacceptable.
- *Acceptable degree: the degree of how acceptable people feel a video clip (1-10)*. It is rated only if the participant thinks the video clip is acceptable.
- *Acceptance grade: the subjective and verbal assessment to the acceptable degree of a video clip*. It has four grades corresponding to the 10 scales of acceptable degree: Fair (1-3), Good (4-6), Very Good (7-8), and Excellent (9-10).

According to the above definitions, the video clips at the acceptance threshold can point to the minimum acceptable combination of encoding parameters. Although unlabeled scales are commonly used in subjective quality assessment, it is not suitable to identify user experience by this method because user experience can only be described in human terminology [38]. Hence, we used not only the 10 scales to differentiate acceptable degrees but also the acceptance grades to illustrate the experience of an acceptable video. To guide participants and make sure they use congruous scores to indicate their experiences, we put the acceptance grades at the positions 1, 4, 7 and 10 under the evaluation slider of acceptable degree (see Fig.1 screenshot 2). Since the acceptance assessments are gained from potential users with real video materials and mobile device under real usage contexts, they in fact contain the influence of user needs, content, device and context, and thereby they are able to represent the user experience of mobile video.

3.1.6. *Data analysis methods*

We collected a total of 50 acceptance records and 10 interview records from the 42 participants. Each acceptance record contains 20 acceptable evaluations to the test clips encoded with 20 combinations of encoding parameters. For the quantitative data (acceptance records) and qualitative data (interview records), we respectively employed statistical methods [39] and the grounded theory [40] to perform the data analysis.

In general, the scores of video quality assessment from a lab experiment are often scanned to eliminate some unreliable data (the filter methods are provided in ITU-R recommendations [8]), regarding the situations that participants may be too optimistic or too pessimistic with a video content, or even sometimes participants vote without taking too much care in watching and tracking the displayed videos. However, in a field user study, the above situations are commonly related to user's preference for the content or the test contexts, which are considered as the important impacts of user experience. Thus, those data should not be easily shifted from data analysis. In this paper, the addressed results are based on the analysis of all the collected data, except the user experience prediction modeling. This is because the user experience prediction models should represent a common user group.

Before performing statistical analysis the normality of the distribution of the acceptance scores was examined by Kolmogorov-Smirnov (K-S) test. The significant level >.05 means a normal distribution, and thus parametric methods can be used for data analysis; otherwise non-parametric methods should be used. In this study, we used ANOVA (Analysis of Variance) (parametric method) and Mann–Whitney *U* test and Wilcoxon signed-rank test (non-parametric method) to compare differences between two or more conditions [39]. Moreover, we used Discriminant analysis [39] to build models of "good" user experience. For understanding the process of data analysis, we explain the statistical methods used in this paper as follows.

One-way ANOVA provides a statistical test of whether or not the means of several groups are all equal. Accompanying with one-way ANOVA, Tukey's HSD test [39] is run to find the significant difference in any comparison of the groups at the alpha level of 0.05. When the sample is exposed to each condition in turn and the measurement of the dependent variable is repeated, a repeated measure ANOVA is more appropriate to be used to compare the difference between different conditions. Since ANOVA is based upon the assumption of the homogeneity of variances, before using ANOVA the homogeneity has to be examined by Levene's test for one-way ANOVA and Mauchly's test of sphericity for repeated measures ANOVA. Without assumption of normality, Mann-Whitney *U* test [39] equals to Student's t-test and can be used to assess the difference between two independent samples; while Mann-Whitney *U* test [39] equals to Student's t-test and can be used to assess the difference between two independent samples; while Wilcoxon test [39] is equivalent to a one-way repeated ANOVA and can be used to measure the difference between two related data sets. In these statistical tests, we adopted the significance level of $p<.05$ to denote that there is a strong evidence of difference.

Discriminant analysis is a technique for classifying a set of observations into predefined classes based on a set of variables known as predictors or input variables. It is employed to gain prediction models of "good" or "not good" user experience based on the encoding parameters. The technique constructs a set of linear functions of the predictors, known as discriminant functions in the form:

$$D = b_1x_1 + b_2x_2 + \ldots + b_nx_n + c$$

Where D is the predicted discriminant score, the X_i is the value of the predictor i, b_i is discriminant coefficients for predictor i, and c is a constant. The accuracy of classification can be computed by comparing the predicted classification with the original classification. This is often well performed with cross-validation. Cross-validation refers to the process of assessing the predictive accuracy of a model in a test sample relative to its predictive accuracy in the learning sample from which the model was developed. Here, using leave-one-out technique, a discriminant function is derived based on all cases except one; the left case is then used as the test sample and reclassified by the discriminant function. The process is repeated for the next case, and so on until each case in the sample is classified by a function. The cross-validated classification provides an estimate of how good the equation would be at classifying new cases.

The qualitative analysis was based on grounded theory [40], an inductive approach, which is developed from the data and moves from the specific to the more general. The method is essentially composed of three elements: concepts, categories and propositions (or "hypotheses"). Whereby concepts are the key elements of analysis since the theory is developed from the conceptualization of data. The derived typical procedures from this theory are: open coding, axial coding and selective coding. Open coding conceptualizes the describing phenomena found in the text; axial coding categorizes the related codes in an advanced stage of development; and selective coding identifies the core category and relates other categories to the core category [40].

With the above methods, all data is analyzed and the results are shown in the next two sections.

4. Quantitative Analysis Results

This section presents the results from the impacts of usage contexts, encoding parameters, and user profiles on the acceptance threshold and the acceptance grade.

4.1. *Content type influence*

Different content types have different minimum acceptable qualities related to the encoding parameters. Fig. 2 shows the percentages of "acceptable" and "unacceptable" for each test clip. For news video, the video quality at 96kbps bit rate can be accepted by most people (80%) regardless of the spatial resolutions and the frame rates (which refer to those used in this paper). For sports video, the video quality at 256kbps can be accepted by over 90% of people. However, the quality of video clips 1, 9, 10, 11 and 12 are below the acceptance threshold (i.e., the unacceptable ratio is over 20%). Concerning the encoding parameters of those video clips (refer to Table 1 in Section 3.1.1), we can see that the sports video clips at 96kbps bit rate are hardly accepted by participants. It can be also observed that at 128kbps video clips are more unacceptable at 25fps than at 12.5fps (clip 9 vs. clip 2, clip12 vs. clip 19), which seems different from the common thought that high frame rate is vital for a fast movement video. This can be explained by the interrelation of encoding parameters that at a given bit rate increasing frame rate will reduce the image clarity, as a result at a low bit rate the blur caused by a high frame rate (25fps) is more annoying than the jitter at 12.5fps [10, 17]. To sum up, the minimum acceptable bit rate is 96kbps for news video and 256kbps for sports video. This conclusion is a little different from that given by Sasse and Knoche [9]. According to the pilot study in [9], on an iPAQ with 320x240 display resolution, at 12.5fps of frame rate, 85% of acceptability can be reached when encoding bit rate is 128kbps for news and 320kbps for football. This may be caused by the different video codecs (WM V8 in [9] and H.264/AVC in this study), the different display devices (iPAQ in [9] and iPhone 3GS in this study), and other different settings. Also, we did not find obvious relations between the spatial resolution and the minimum acceptable quality for both content types. However, it was presented that the acceptability decreases as the image size reduces for sports video in [9]. The inconsistency may be due to the difference of used image sizes in the two studies (maximum 240x180 pixels is used in [9]).

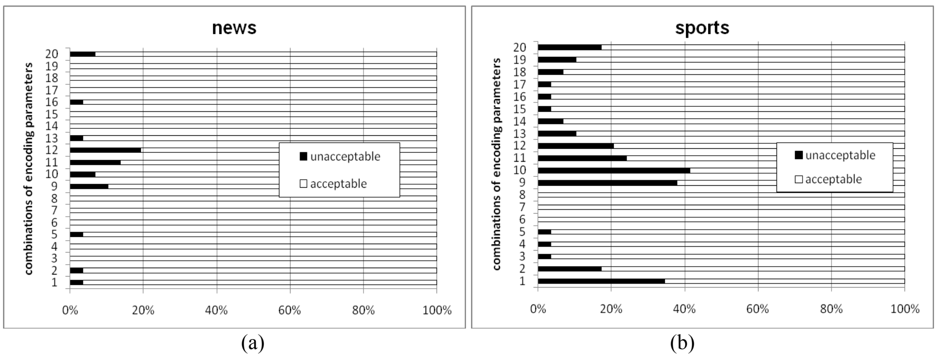
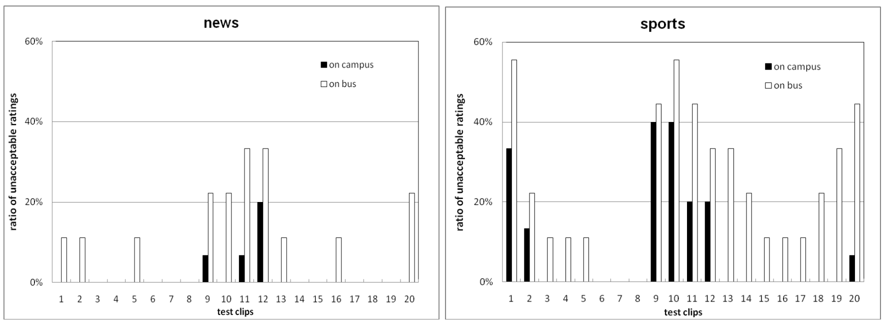


Fig. 2. Acceptability of different encoding combinations

To examine if the content type impacts the acceptable degree, we ran a set of Mann-Whitney U tests for 20 pairs of video clips of the news and sports videos with the mean scores of acceptable degree as the dependent variables. Significant difference was only found at 5 pairs of video clips with the number of 1, 2, 3, 9 and 10. The corresponding encoding parameters are 96kbps and128kbps (or 256kbps) at the resolution of 480x320 pixels ($p<.05$). This situation is expected because of the higher unacceptable ratio of sports video than news video at low bit rate. However, when resolution is 320x240 pixels or when resolution is 480x320 and the bit rate greater than 256kbps, there is no significant difference between the two content types ($p>.05$). Thus, it can be concluded that at a small image resolution or once mobile video quality exceeds the acceptance threshold, users are unable to tell the difference of acceptable degree between various content types encoded with same encoding parameters.

4.2. *Context influence*

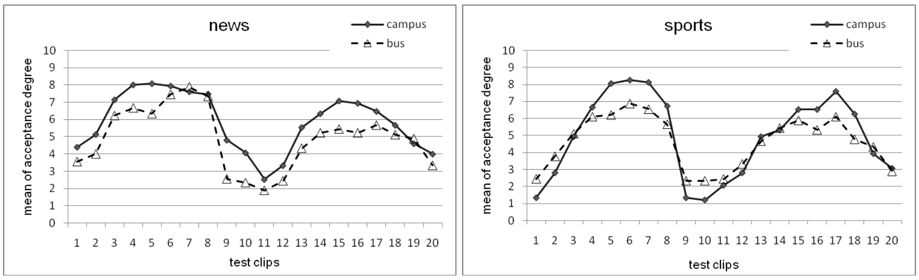
Apart from the influence of content types, usage contexts also influence the acceptability. The most unacceptable ratings were given in the context of "bus" (81% for news and 62% for sports) rather than "campus" (19% for news and 38% for sports). Specifically, the unacceptable ratings under the campus context only centered on the low bit rate video clips; whilst under the bus context some viewers treated the clips at high bit rate as unacceptable (Fig. 3). This reveals that whether people accept a mobile video content is to some extent influenced by the context.



(a) (b)

Fig. 3. Impacts of contexts on unacceptable ratio

For each content type, one-way ANOVA was conducted with the scores of acceptable degree as dependent variables and the contexts of "campus" and "bus" as the factors. The result indicated that the significant impact of context on the acceptance degree mainly existed at high bit rate ($p<.05$). Fig. 4 shows the curves of the mean scores of acceptable degree for news and sports video.



(a) (b)
Fig. 4. Imapcts of contexts on acceptable degree

4.3. *Encoding parameters influence*

The relationships of encoding parameters with the acceptance threshold have been discussed with the effects of content types together in Section 4.1. Thereby, we only declare how the encoding parameters influence the acceptable degree. Repeated measures ANOVA or Wilcoxon test was used for the following data analysis.

Effects of bit rate – It is significant that the acceptable degree increases with bit rate at a same condition of spatial resolution and frame rate ($p<.05$). However the overall correlation is not linear but logarithmic (Fig. 5). Based on one-way ANOVA, there is strong evidence that the mean scores of acceptable degree are not different at the bit rate of between 96kbps and 128kbps ($p=.851$), 512kbps and 768kbps ($p=.990$); and the difference is not significant between the adjacent bitrates 256kbps and 512kbps ($p=.323$). However, there is significant difference between 128kbps and 256kbps (MD (mean difference) = 2.44, $p<.001$). Moreover, the situation was found for both content types ($p=.014$ and $p=.005$ for news and sports respectively). It indicates that people's perception to video quality will be markedly improved when the encoding bit rate becomes over 256kbps; and the perception will not significantly increase when the encoding bit rate is over 512kbps.

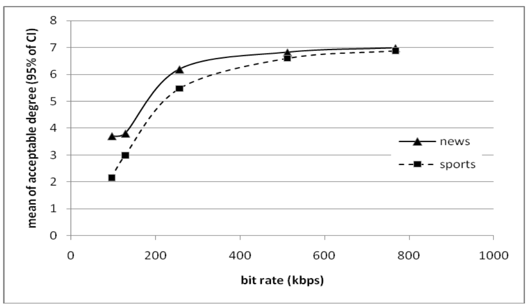


Fig. 5. Impact of bit rate on acceptable degree

Effect of spatial resolution - The effect of spatial resolution on the acceptable degree is statistically significant when the bit rate is equal and greater than 256kbps for news video (Wilcoxon test, $p<.05$). For sport video, however, the effect of spatial resolution is dependent on not only bit rate but also frame rate. At 12.5fps, the significant difference of acceptable degree between two spatial resolutions is only found at the lowest bit rate 96kbps and the highest bit rate 768kbps; while at 25fps, the significant difference happens at all bit rates except 96kbps ($p<.05$). Fig. 6(a) and Fig. 6(b) compare the means of acceptable degree at two spatial resolutions (480x320 and 320x240). It can be seen that for news participants always prefer video a big resolution; while for sports video, the small spatial resolution (320x240) is preferred when bit rate is low.

Effect of frame rate - Under a given spatial resolution, whether high or low frame rate is more acceptable rests with bit rate. When spatial resolution is 320x240, video clips at 25fps are more acceptable than those at 12.5fps if bit rate is equal and greater than 512kbps; on the contrary, 12.5fps takes precedence of 25fps if bit rate is less than 512kbps. While when spatial resolution is 480x320, the division is at 256kbps. This trend does not vary with the content types (Fig. 6(c), Fig. 6(d)). However, the difference of the acceptable degrees between the two frame rates is more significant for sports video than news video. It indicates the sensitivity of videos with fast motion to frame rate.

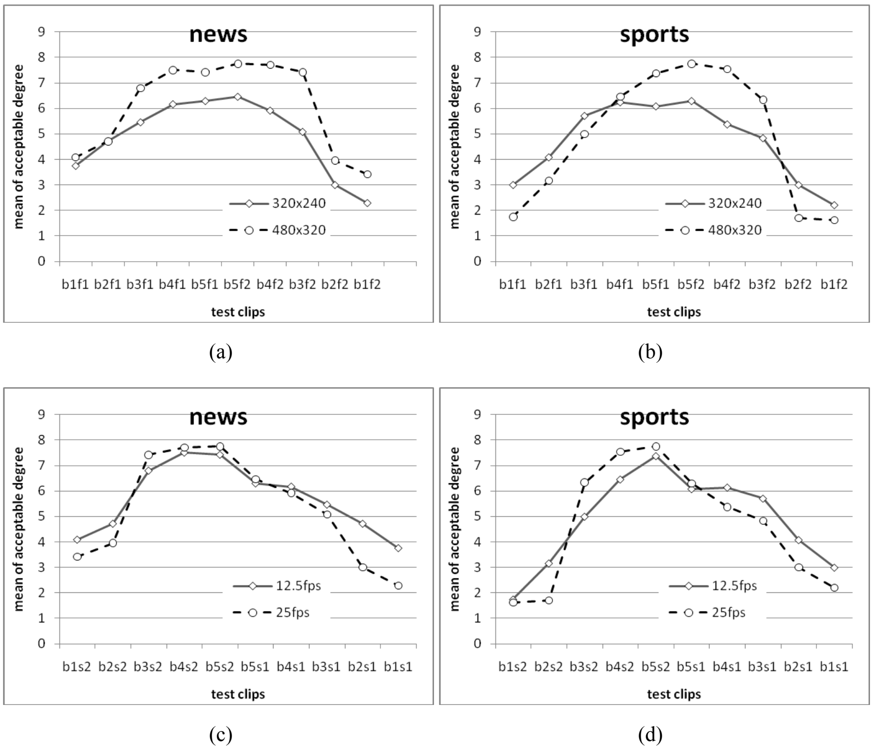


Fig. 6. Impacts of spatial resolutions and frame rates on acceptable degree. (a) (b) means of acceptable degrees at spatial resolutions 320x240 and 480x320 for news video and sports video; (c) (d) means of acceptable degrees at frame rates 12.5fps and 25fps for news video and sports video

4.4. *User profiles influence*

User profile is the information to define the characteristics of a user. We collected four aspects of user profile; age, gender, preference for video content, experience in watching mobile video. To test the effect of user profiles on users' acceptance degree to mobile video quality, we conducted a set of Wilcoxon tests. The analysis results are described as follows.

Age has a significant impact on acceptable degree at high bit rate. To test the effect of age on users' acceptance degree to mobile video quality, we separated participants into three age groups: <=21 (13 people gave14 records), 22–29 (19 people gave 25 records), and >=30 (10 people gave 11 records). According to the Mann-Whitney test, there is no difference of the acceptance degree between three age groups at low bit rate such as 96kbps and 128kbps (p>0.25); however, significant difference is found at high bit rate (512kbps and 768kbps) (p<0.05). The details were observed that younger people (the age groups of <=21 and 22–29) gave a much higher acceptance score to the videos with high bit rate and high spatial resolution (480x320) than the elder people (the age group of >=30) and the mean difference between the younger groups and the elder group was more than 1.9; whereas if the spatial resolution is 320x240, only the age group of 22–29 gave significant higher scores to the video at 256kbps to 768kbps than the group of over 30 (p<.05).

Preference for video content has a significant impact on acceptable degree for sports video (p<.05). The more people like the sport video, the higher their acceptable degree is (Fig. 7). However, this does not happen with news video. Probably it is due to the fact that most participants (59%) have a neutral attitude to news video.

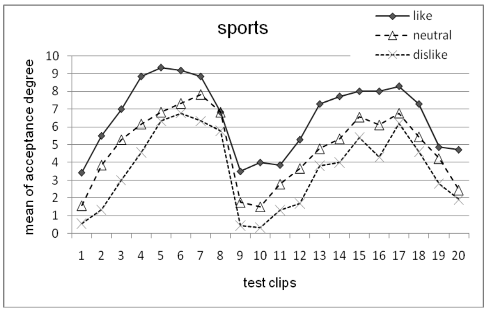
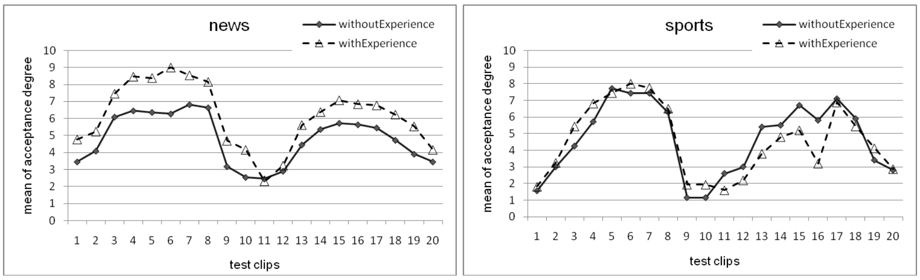


Fig. 7. Impact of pereference for video content for sports video on acceptable degree

In contrast with the people's preference, **people's experience in mobile video** only impacts the acceptable degree for news video (p<.05). The significant difference is found at high bit rate (768kbps and 512kbps) and high spatial resolution (480x320). From Fig. 8, we can see the people who have had experience in viewing videos on mobile phone give much higher evaluation than the people without the experience.

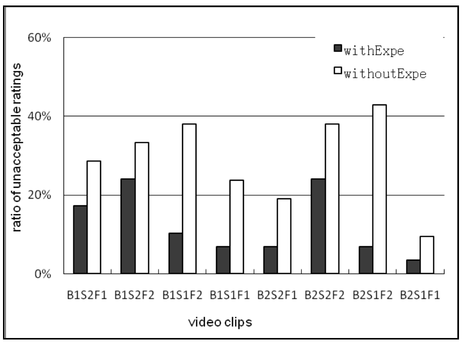


(a) (b)

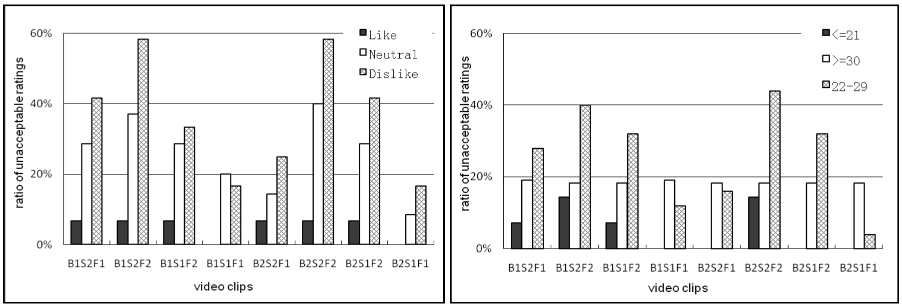
Fig. 8. Impacts of experience in mobile video on acceptable degree

Gender has no impact on acceptable degree.

According to the analysis of section 4.1, the minimum bit rate of acceptable video quality (based on 80% acceptable ratio or 20% unacceptable ratio of the total ratings) is generally 96kbps for news video and 256kbps for sports video. However, the threshold seems different for diverse user groups. The unacceptable ratios for different user groups are illustrated in Fig 9. From Fig. 9(a), it can be observed that people without experience in viewing mobile video are more ready to reject the video quality at low bit rates (96kbps and 128kbps), while the people with experience only reject the quality at low bit rates and high spatial and temporal resolutions 480x320&25fps. Fig. 9(b) shows that people's preference for content types significantly influences their minimum acceptable video quality. If people are interested in the video content, they can accept a very low quality (such as 96kbps, 480x320, 25fps); while if they do not really like the content, they can only accept the low bit rate (96kbps) video at 320x240 & 12.5fps. Fig. 9(c) presents the impact of age groups on the unacceptable ratio. The youngest group (age<=21) is most easily accepting the low video quality; yet the middle young group (age range of 22–29) is the hardest group to be satisfied. Their minimum acceptable video quality is that encoded with 320x240 & 12.5fps at 96kbps or 12.5fps at 128kbps.



(a)



(b) (c)

Fig. 9. Impacts of user profiles on unacceptable ratio. (a) experience in mobile video; (b) preference for video content; (c) age group

5. Qualitative Analysis Results

As the statistical analysis presented, the acceptance evaluation differs from person to person, context to context. To understand how users evaluate mobile video's quality and to explain why there are obvious discriminations between user experiences, this section discusses the analysis results of qualitative data (interview transcript and field notes) from two aspects: users' evaluation criteria and context impact on user experience.

5.1. *Users' evaluation criteria*

About people's evaluation criteria, we designed the following structure of questions.
Main Questions:

How did you assess/distinguish the video's acceptable degree?
What criteria did you use to define the difference between acceptable levels?
What are the differences between the acceptance grades in your opinion?

Supporting Questions:

What were the main factors you used to evaluate the video quality?
What is your expectation of an excellent video, or in your opinions what an unacceptable video looks like? (If someone did not rate any video as a particular scale, e.g., excellent or unacceptable).

Based on the responders' answers, we extracted the main concepts and categorized them, then summarized the evaluation criteria as follows.

- **Excellent**: high clarity, smooth movement, clear facial details, attractive content
- **Very Good**: no blur at all, smooth movement, clear facial features
- **Good**: clear face, mild fuzzy, easy to follow the ball, willing to watch it regularly and watch for a long time
- **Fair**: being able to know the information, recognizable faces, fuzzy, crispy, unwilling to watch for a long time,
- **Unacceptable**: obvious block effect, indiscernible objects, frame jump, very uncomfortable to watch, completely not interesting.

Considering the characteristics of each grade, four main conclusions can be drawn.

(i) Block effect has the worst impact on user experience, followed by blur and frame jump (low frame rate).

(ii) Whether or not a face can be recognized is very important for the videos mainly about people; while whether the game and the ball can be followed is very important for the sports video. This confirms the qualitative comments in [10].

(iii) Interest exerts a significant impact when distinguishing if a video quality is unacceptable and if an excellent user experience can be achieved.

(iv) There is an important discrepancy between the acceptance grades of *Fair* and *Good*, that is, users will frequently watch the mobile videos only if they feel the quality reaches good.

5.2. *Context impact on user experience*

Most respondents selected from the campus did not think the context really affected their watching even if environment was noisy and crowded. Only one person mentioned that the sunlight might result in her assessment outdoor differing from indoor. On the bus, however, some people felt sick when they were watching videos; and some felt difficult to concentrate on the video content with a low quality. Moreover, we observed that during the short test time (10 minutes), participants on the campus only gave a glance to the surroundings for a few times, but participants on the bus were absent-minded and frequently draw attention out of the video, especially when the bus stopped due to traffic or bus stops. This actually leads to a big deviation of acceptable degree scores on the bus.

5.3. *Coherency of users' opinions with the results of qualitative analysis*

When comparing user's opinions with the results of qualitative analysis, a high consistency can be found in several respects. Firstly, the distortion of block effect is mainly caused by coarse quantization. Under a given bit rate, the quantization becomes coarser with the increase of the spatial resolution and the frame rate. This is why the video clips at low bit rate and 480x320&25fps are more unacceptable than others. Secondly, interest or preference takes an action in people's judgments. Thirdly, usage context impacts user's viewing experience via obstructing user's attention to the video.

Participant's words may well explain the phenomena that different user groups have different acceptance grades and acceptance threshold. The people without any experience of viewing videos on mobile phone have a high expectation of the mobile video's quality. In their opinions, *the quality of mobile video should give them as good experience as videos displaying on a PC*. But the teenager said that *all videos looked good and were better than YouTube videos*. In respect of the preference for the content, a participant said: "*If I like the content, I almost accept all of them only if the quality is too bad because I enjoy the content and like to watch it continually. But if I don't like the content, e.g., sports, I am fussier to judge the video quality because my focus is not the content but its quality.*"

6. Strategies for User Experience Optimization

Since different user groups have different perceptions and expectations to mobile video quality, it is necessary to consider mobile video delivery from a user's perspective. In terms of the analysis in Section 4 and 5, we have already obtained some knowledge on user's experience on mobile video, such as how user experience is influenced by user profiles and encoding parameters. According to these findings, we propose two strategies to manage the delivery of mobile video. One is user-driven acceptance threshold adaptation, which aims to support an acceptable video quality to different user groups under the circumstance that the video bit rate/transmission bandwidth is restricted. The other is in a form of user experience prediction models, which can be used to predict if a provided video quality can achieve a good user experience.

6.1. *User–driven acceptance threshold adaptation*

The user-driven acceptance threshold adaptive strategy is inspired by the fact that the minimum acceptable mobile video quality is different from content to content, and user to user, thus the differences should be taken into account for providing an acceptable video quality to users under the condition of limited bit rate/bandwidth. The proposed adaptation first checks if the user information is available. If there is no user profile information, it will manage the acceptable quality according to the content type; otherwise, it will manage the video quality based on the existing user information, including experience on mobile video, preference to content, and age group. The pseudo code of the adaption is documented in Appendix A.

This adaptation strategy is feasible for the reasons: (i) video content types can be known beforehand; and (ii) user information can be assigned by users themselves and stored into their mobile terminals; and it can be submitted to the Server by terminal programs when requesting the video service.

6.2. *User experience prediction*

Since whether users have a "Good" viewing experience determines whether or not they are willing to watch the mobile videos in the long term, it is vital to predict if a provided video quality can achieve a good user experience. Therefore, we employed discriminant analysis to classify the video quality into two groups: *good* and *not good* on the basis of different encoding parameter combinations. The analysis steps are the following.

Firstly, to avoid the interference of arbitrary ratings or too positive and negative ratings, we scanned all the collected acceptable scores and removed 6 records, of which, 5 records were too positive or negative ratings (all clips were marked >=8 or <=1) and 1 record was obviously disordered.

Secondly, with the acceptance grade "Good" as the boundary, we transformed the interval variable of the acceptable score into a nominal variable –"IsGood", in which the "good" cases are those that gain the mean score of acceptable degree greater than and equal to 4, and the "not good" are the rest.

Thirdly, we used the frame rate and the transformed bit rate and spatial resolution as the predictors. Due to the non-linearity of the influence of bit rate on the acceptance, the value of bit rate cannot be directly used for a linear prediction. Therefore, we transformed the bit rate as natural logarithm bit rate (LBR): 4.56, 4.85, 5.55, 6.24, and 6.64

respectively corresponding to 96kbps to 768kbps. In addition, the big value of spatial resolution will lead to a very small coefficient if it is used directly. Thereby we used the transformed spatial resolution, so-called pixels per mm^2 (PPM), to represent it. Since the screen area of iPhone is 75mm x 50mm (4,125mm^2), the average pixels per mm^2 is 37pixels for the spatial resolution of 480x320 (153,600 pixels) and 19 pixels for 320x480 (76,800pixels).

Finally, the discriminant analysis using stepwise method was conducted for each content type and each usage context. The stepwise discriminant analysis can produce the smallest set of independent variables (LBR, PPM, and FR) which correctly classify the largest number of cases. The derived classification functions are shown as follows, whereby D is the discriminant score.

Relaxed scenario (on campus):

(i) News video

$$D_{good} = 11.783\times LBR + 0.479\times PPM + 0.329\times FR - 44.292$$
$$D_{not\ good} = 9.591\times LBR + 0.378\times PPM + 0.439\times FR - 32.413$$

(ii) Sports video

$$D_{good} = 14.567\times LBR + 0.262\times PPM - 47.34$$
$$D_{not\ good} = 11.984\times LBR + 0.299\times PPM - 34.414$$

Nervous scenario (on bus):

(i) News video

$$D_{good} = 13.001\times LBR + 0.303\times FR - 41.261$$
$$D_{not\ good} = 10.543\times LBR + 0.401\times FR - 30.161$$

(ii) Sports video

$$D_{good} = 16.530\times LBR - 50.395$$
$$D_{not\ good} = 13.653\times LBR - 34.600$$

From the above functions, we can see that not all encoding parameters contribute to the discrimination for various situations. For example, in the scenario of viewing sports videos on a bus, only bit rate determines if user's experience is good.

The Wilks' lambda [39] was significant for the four situations (Λ=0.715, 0.629, 0.623, 0.572; $p<.001$), which indicates that the models including the above variables were able to significantly discriminate the two groups. For a given combination of encoding parameters (case), the both classification scores (D_{good} and $D_{not\ good}$) are computed and compared. And the case will belong to the group with higher classification score. To verify the predictive accuracy of the classification models, cross-validation procedure was performed for each content type and context. Table 2 lists the percentage of correct prediction. Taking the sports video viewed on campus as an example, the table demonstrates that for the total sample of 280 cases, 233 (83.2%) overall are classified correctly. Correct classification rates of 82.7% are observed for "Good" group and 84.2% for "Not good" group.

Table 2. Accurate percentage of classification prediction

Group	News & Campus	Sports &Campus	News & Bus	Sports & Bus

		(C=236,N=300)		(C=233,N=280)		(C=111,N=140)		(C=113,N=140)	
		Good	Not good	Good	Not good	Good	Not good	Good	Not good
%	Good	75.3	24.7	82.7	17.3	76.2	23.8	85.5	14.5
	Not good	6.8	93.2	15.8	84.2	11.4	88.6	25	75

These prediction models can be used to estimate if a given encoding parameters for a given video content can achieve a good user experience. Furthermore, it can be used to control the mobile video streaming for achieving a good user experience by adjusting the encoding parameters.

7. Conclusion

In order to understand the user experience of viewing mobile video, a field user study was conducted to find out what is the minimum acceptable quality of mobile video and how user's acceptance of mobile video is influenced by content type, usage context, encoding parameters, and user profiles. The main findings can be summarized into four respects.

First, consistent with the conclusion in [9], the minimum acceptable bit rate varies from content to content. In this paper, with 80% of the acceptable ratings as the acceptance threshold, the minimum acceptable bit rate is 96kbps for news video and 256kbps for sports video. However, above the acceptance threshold the acceptable degree for different content types is not significant different under the same encoding condition.

Second, people are more likely to accept a low quality video under a relaxed context (such as on campus) than a nervous usage context (such as on a bus).. For high quality videos, the acceptable degree is much higher on the relaxed context than on the nervous context.

Third, the encoding parameters have impacts on the acceptable degree. When the bit rate becomes over 256kbps mobile user's experience dramatically increases; and it keeps steady when the bit rate is over 512kbps. For news video, users always prefer a big image size (480x320); but for sports video, when bit rate decreases a small spatial resolution (320x240) is preferred. There are two division points of distinguishing which frame rate is preferred. Given the image size of 320x240 pixels, if the bit rate is below 512kbps, the more acceptable frame rate is 12.5fps; otherwise it is 25fps. Given the image size of 480x320 pixels, the division bit rate is at 256kbps. Such tendency is more obvious for videos with large amounts of motion (e.g., sports videos).

Fourth, user profiles, except gender, influence user's acceptance to mobile video. For users with different ages, younger people (<30) often give a much higher acceptance score to the high bit rate video than older people (>=30). However, young people have a different attitude to the low quality video: people at the age of below 22 are very likely to accept a low video quality; while people at age range of 22–29 highly reject it. User's preference for video content significantly impacts the acceptable degree for sports video; conversely, whether or not a user has experience of mobile video only impacts the acceptable degree for news video. Regarding the acceptance threshold, users' demand for the acceptable quality decreases with the increase of their interest in a video content but increases with their experience of viewing mobile video.

All the above findings are well supported by the qualitative study. Furthermore, through analyzing the interview data, we found that it is important to determine not only whether a mobile video is acceptable but also whether a good user experience can be achieved. The acceptable boundary is influenced by content type, context and user profile, and the good experience decides whether users are willing to use mobile video service in the long term. Therefore, we propose two strategies for optimizing the user experience. One is the user-driven acceptance threshold adaptation, which focuses on a limited bit rate condition and aims to provide an acceptable video quality to mobile users based on their diverse requirements. The other is the user experience prediction models, which provides a prediction of good user experience based on a set of encoding parameters.

The limitations of this study are mainly about limited content types and usage contexts. It should be noticed that the results might still be influenced by the used specific video contents although we selected common materials to represent news and sports videos. It should be also pointed out that due to the limited respondents, the interaction effects of user profiles are not studied in this paper. Our further study will involve more typical mobile video contents and the usage contexts so that more accurate and flexible user experience strategies can be built and applied into real applications. In addition, the research approach of combining qualitative and quantitative into the user experience optimization strategy is tentative. It can be improved in the future.

Appendix A. Pseudocode of User-driven Acceptance Threshold Adaptation

```
IF bit rate limitation is greater than 128kbps THEN
   encode video at a bit rate > 128kbps
ELSE
   IF unknown user profile THEN
      // determine the encoding parameters according to content types
      // in this paper, only two content type are studied
      IF news video THEN
         encode the video at a bit rate >= 96kbps
      ENDIF
      IF sports video THEN
         IF bit rate limitation is equal to 128kbps THEN
            encode video at the given bit rate and frame rate 12.5fps
         ELSE
            encode video at a bit rate>=96kbps, spatial resolution 320x240, and frame rate 12.5fps
         ENDIF
      ENDIF
   ELSE  // some user information is known
      IF a favorite content type THEN
         encode the video at a bit rate >= 96kbps
      ELSE // not favorite content type
         IF a new user THEN
            encode the video at the bit rate 128kbps,
            spatial resolution 320x240, and frame rate 12.5fps
         ELSE //a mature user
            encode the video at the bit rate>=96kbps, any combinations of spatial resolutions
            and frame rates except concurrent 480x320&25fps
         ENDIF
         IF Age is between 22 and 29 THEN
            IF bit rate limitation is equal to 128kbps THEN
               encode video at the given bit rate and frame rate 12.5fps
            ELSE
               encode video at a bit rate>=96kbps, spatial resolution 320x240, and frame rate 12.5fps
            ENDIF
         ELSE   // other ages
            encode the video at a bit rate >= 96kbps
         ENDIF
      ENDIF
   ENDIF
ENDIF
```

References

[1] M. Hassenzahl and N. Tractinsky, "User experience - A research agenda," *Behaviour & Information Technology,* vol. 25, pp. 91-97, 2006.

[2] S. Jumisko-Pyykkö and J. Häkkinen, "Evaluation of subjective video quality on mobile devices," in *13th ACM International Conference on Multimedia*, Singapore, 2005, pp. 535-538.

[3] M. Ries, O. Nemethova, and M. Rupp, "Video Quality Estimation for Mobile H.264/AVC Video Streaming," *COMMUNICATION,* vol. 3, pp. 41-50, Janurary 2008.

[4] G. Zhai, J. Cai, W. Lin, X. Yang, W. Zhang, and M. Etoh, "Cross-dimensional perceptual quality assessment for low bit-rate videos," *IEEE Transactions on Multimedia,* vol. 10, pp. 1316-1324, November 2008 2008.

[5] H. Knoche, J. D. McCarthy, and M. A. Sasse, "Can small be beautiful?: assessing image resolution requirements for mobile TV," in *13th annual ACM international conference on Multimedia*, Hilton, Singapore, 2005, pp. 829-838.

[6] H. Knoche, J. D. McCarthy, and M. A. Sasse, "How low can you go? The effect of low resolutions on shot types in mobile TV," *Multimedia Tools and Applications,* vol. 36, pp. 145-166, 2008.

[7] S. Jumisko-Pyykkö, V. K. M. Vadakital, and M. M. Hannuksela, "Acceptance threshold: A bidimensional research method for user-oriented quality evaluation studies," *International Journal of Digital Multimedia Broadcasting,* vol. 2008, p. 20, 2008.

[8] ITU-R Recommendation BT. 500-11:, "Methodology for the Subjective Assessment of Quality for Television Pictures," 2004.

[9] M. A. Sasse and H. Knoche, "Quality in Context-an ecological approach to assessing QoS for mobile TV," in *2nd ISCA/DEGA Tutorial & Research Workshop on Perceptual Quality of System,* Berlin, Germany, 2006.

[10] J. D. McCarthy, M. A. Sasse, and D. Miras, "Sharp or smooth?: comparing the effects of quantization vs. frame rate for streamed video," in *SIGCHI Conference on Human Factors in Computing Systems*, San Francisco, California, United States, 2004, pp. 535-542.

[11] S. Orgad, "This box was made for walking," Department of Media and Communications, London School of Economics and Political Science 2006.

[12] I. D. Constantiou, "Consumer behaviour in the mobile telecommunications' market: The individual's adoption decision of innovative services," *Telematics and Informatics,* vol. 26, pp. 270-281, 2009.

[13] K. Miyauchi, T. Sugahara, and H. Oda, "Relax or study?: A qualitative user study on the usage of Mobile TV and video," in *6th EuroITV*. vol. 5066 Salzburg, Austria: Springer, 2008, pp. 128-132.

[14] P. Repo, K. Hyvönen, M. Pantzar, and P. Timonen, "Users inventing ways to enjoy new mobile services-the case of watching mobile videos," in *37th Hawaii International Conference on System Sciences (HICSS)*, Hawaii, USA, 2004.

[15] H. Knoche and J. D. McCarthy, "Mobile users' needs and expectations of future multimedia services," in *Wireless World Research Forum (WWRF)12,* Toronto, Canada, 2004.

[16] S. Jumisko-Pyykkö and M. M. Hannuksela, "Does context matter in quality evaluation of mobile television?," in *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services.* Amsterdam, The Netherlands: ACM, 2008.

[17] W. Song, D. Tjondronegoro, and S. Azad, "User-centered video quality assessment for scalable video coding of H.264/AVC standard," in *16th International Multimedia Modeling Conference,* Chongqing, China: Springer, 2010, pp. 55-65.
[18] S. Jumisko-Pyykkö, U. Reiter, and C. Weigel, "Produced quality is not perceived quality - A qualitative approach to overall audiovisual quality," in *3DTV Conference, 2007*, 2007, pp. 1-4.
[19] F. Agboma and A. Liotta, "User centric assessment of mobile contents delivery," in *4th International Conferences on Advances in Mobile Computing and Multimedia,* Yogyakarta, Indonesia, 2006, pp. 121-130.
[20] D. Wang, F. Speranza, A. Vincent, T. Martin, and P. Blanchfield, "Toward optimal rate control: a study of the impact of spatial resolution, frame rate, and quantization on subjective video quality and bit rate," in *Visual Communications and Image Processing 2003*, 2003, pp. 198-209.
[21] L. Sahafi, T. S. Randhawa, and R. H. S. Hardy, "Context-based complexity reduction of H.264 in video over wireless applications," in *6th Workshop on Multimedia Signal Processing*, 2004, pp. 23-26.
[22] M. Claypool and A. Tripathi, "Adaptive video streaming using content-aware media scaling " Worcester Polytechnic Institute, Worcester, Massachusetts, USA, Computer Science Technical Report Janurary 7, 2004 2004.
[23] Z. Lu, W. Lin, E. Ong, S. Yao, S. Wu, B. C. Seng, and S. Kato, "Content-Based Quality Evaluation on Frame-Dropped and Blurred Video," in *Multimedia and Expo, 2007 IEEE International Conference on*, 2007, pp. 1455-1458.
[24] Y. Cui, J. Chipchase, and Y. Jung, "Personal TV: A Qualitative Study of Mobile TV Users," in *EuroITV 2007*, P. C. e. al., Ed.: Springer, 2006, pp. 195-204.
[25] M. Vangenck, A. Jacobs, B. Lievens, E. Vanhengel, and J. Pierson, "Does mobile television challenge the dimension of viewing television? An explorative research on time, place and social context of the use of mobile television content," in *EuroITV 2008*, Salzburg, Austria, 2008, pp. 122-127.
[26] H. Knoche and J. D. McCarthy, "Design requirements for mobile TV," in *Mobile HCI'05*, Salzburg, Austria, 2005, pp. 69-76.
[27] K. O'Hara, A. S. Mitchell, and A. Vorbau, "Consuming video on mobile devices," in *SIGCHI on Human Factors in Computing Systems*, San Jose, CA, USA, 2007, pp. 857-866.
[28] C. Södergård, "Mobile television-technology and user experiences," VTT Information Technology 506, 2003.
[29] J. Chipchase, C. Yanqing, and Y. Jung, "Personal television: A qualitative study of mobile TV users in South Korea," in *Mobile HCI*, Espoo, Finland, 2006.
[30] S. Buchinger, S. Kriglstein, and H. Hlavacs, "A comprehensive view on user studies: Survey and open issues for mobile TV," in *7th European Conference on Interactive Television*, Leuven, Belgium, 2009, pp. 179-188.
[31] S. Jumisko-Pyykkö, M. Weitzel, and D. Strohmeier, "Designing for user experience: what to expect from mobile 3D TV and video?," in *1st international conference on Designing interactive user experiences for TV and video,* vol. 291 Silicon Valley, California, USA ACM, 2008, pp. 183-192.
[32] S. Jumisko-Pyykkö, M. Weitzel, and D. Strohmeier, "Designing for user experience: what to expect from mobile 3D TV and video?," in *1st International Conference on Designing*

Interactive User Experiences for TV and Video, Silicon Valley, California, USA 2008, pp. 183-192.
[33] L. Eronen, "User centered research for interactive television," in *2003 European Conference on Interactive Television,* Brighton, UK, 2003, pp. 5-12.
[34] R. Bernhaupt, M. Obrist, and M. Tscheligi, "Usability and usage of iTV services: Lessons learned in an Austrian field trial," *ACM Computers in Entertainment,* vol. 5, August 2007.
[35] ITU-T, "Subjective Video Quality Assessment Methods for Multimedia Applications," P.910 Recommendation, 1999.
[36] G. Ghinea and J. P. Thomas, "Quality of perception: user quality of service in multimedia presentations," *IEEE Transactions on Multimedia,* vol. 7, pp. 786-789, 2005.
[37] C. Carlsson and P. Walden, "Mobile TV-to live or die by content," in *40th Annual Hawaii International Conference on System Sciences (HICSS)*, 2007, pp. 51-57.
[38] Nokia. *Quality of experience (QoE) of mobile services: Can it be measured and improved?* 2006 [cited 30 Aug 2009]. Available from www.nokia.com/NOKIA_COM_1/ Operators/Downloads/Nokia_Services/whitepaper_qoe_net.pdf.
[39] A. Aron, E. Aron, and E. J. Coups, *Statistics for psychology*: Harlow: Pearson Education, 2009.
[40] A. Strauss and J. Corbin, *Basic of qualitative research: Grounded theory procedures and techniques*: Sage Publications, 1990.